Lecture 4
Single View Metrology

Professor Silvio Savarese
Computational Vision and Geometry Lab

Lecture 4 - Silvio Savarese 16-Jan-15

• Review calibration and 2D transformations
• Vanishing points and lines
• Estimating geometry from a single image
• Extensions

Reading:
[HZ] Chapter 2 “Projective Geometry and Transformation in 2D”
[HZ] Chapter 3 “Projective Geometry and Transformation in 3D”
[HC] Chapter 8 “More Single View Geometry”
[Hoiem & Savarese] Chapter 2

Calibration Problem

The calibration problem was discussed in details during lecture 3.

\[ P \rightarrow M \rightarrow \mathbf{p}_i = \begin{bmatrix} u_i \\ v_i \end{bmatrix} \]

\[ M = K[R \ T] \]

\[ K = \begin{bmatrix} f & -f\cot\theta & u_c \\ 0 & 1 & v_c \\ 0 & 0 & 1 \end{bmatrix} \]
Once the camera is calibrated (intrinsics are known) and the transformation from the world reference system to the camera reference system (which accounts for the extrinsics) is also known, can we estimate the location of a point P in 3D from its observation p? The general answer to this question is no. This is because, given the observation p, even when the camera intrinsics and extrinsic are known, the only thing we can say is that the point P is located somewhere along the line defined by C and p. This line is called the line of sight.

The actual location of P along this line is unknown and cannot be determined from the observation p alone.
Recovering structure from a single view

In the remainder of this lecture, we will introduce tools and techniques for inferring the geometry of the camera and the 3D environment from a just one image of such environment.

Transformation in 2D

- Isometries
- Similarities
- Affinity
- Projective

Before we go into the details of that, let me recap some of the concepts you have already explored when we talked about transformations. There are four important transformations in 2D. All these transformations will be described in homogenous coordinates.

The first one is called isometric transformation which is in general the concatenation of a rotation and translation transformation and expressed by the matrix $H_e$ and Eq. 4. This transformation preserves the distance between any pair of points and has 3 degrees of freedom (2 for translation, 1 for rotation). This group captures the motion of a rigid object.

Isometries: [Euclidean]

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = H_e \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad [\text{Eq. 4}]$$

- Preserve distance (areas)
- 3 DOF
- Regulate motion of rigid object
Transformation in 2D

**Similarities:**

\[
\begin{bmatrix}
    x' \\
    y' \\
    1
\end{bmatrix}
= \begin{bmatrix}
    SR & t \\
    0 & 1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix}
= H_s
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix}
\]

\[
S = \begin{bmatrix}
    s & 0 \\
    0 & s
\end{bmatrix}
\]

- Preserve
  - ratio of lengths
  - angles
- 4 DOF

**Affinities:**

\[
\begin{bmatrix}
    x' \\
    y' \\
    1
\end{bmatrix}
= \begin{bmatrix}
    A & t \\
    0 & 1
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix}
= H_s
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix}
\]

\[
A = \begin{bmatrix}
    a_{11} & a_{12} \\
    a_{21} & a_{22}
\end{bmatrix}
= R(\theta) \cdot R(-\phi) \cdot D \cdot R(\phi)
\]

\[
D = \begin{bmatrix}
    s_x & 0 \\
    0 & s_y
\end{bmatrix}
\]

- Preserve:
  - Parallel lines
  - Ratio of areas
  - Ratio of lengths on collinear lines
  - others...
- 6 DOF

The second group of transformations is called similarity transformation. The transformation is a concatenation of a translation, rotation and scale translation transformations. It preserves the ratio of lengths between any two line segments before and after transformation. It also preserves the angle between any intersecting lines. The group has 4 degrees of freedom (2 translation, 1 rotation, 1 scale).

The third transformation group is the affine transformation. This group can be interpreted as a series of translation, rotation and anisotropic scaling transformations. The core of the affine transformation is characterized by the matrix A which includes a rotation by phi, an anisotropic scaling by [sx 0; 0 sy], an inverse rotation by –phi and an arbitrary rotation by theta (see Eq. 7). This transformation has also a mathematical interpretation: Any positive definite 2x2 matrix can be decomposed into two orthogonal matrices and a diagonal matrix by Singular Value Decomposition (this is by assuming positive scaling; for negative scaling, the transformation may not be unique.) Thus we can express A = UDV = (UV^T) (VDV^T) = R(theta) R(-phi) D R(\phi) by replacing UV^T = R(theta), V = R(\phi)^T, and where D is a diagonal matrix with the singular values. Note that transpose of a rotation matrix is the inverse of the rotation matrix. Thus, an arbitrary 2x2 matrix can be decomposed into R(theta), R(\phi), D.

This transformation group preserves parallel lines (parallel lines are still parallel after transformation). The ratio of lengths on collinear lines is preserved after transformation and thus it follows that the ratio of areas within any arbitrary shapes is preserved.

The transformation has 6 degree of freedom (4 elements in the matrix A, 2 for translation).
The cross ratio

The cross-ratio of 4 collinear points

\[ \frac{||P_3 - P_1||}{||P_3 - P_2||} \cdot \frac{||P_4 - P_2||}{||P_4 - P_1||} \]

[Eq. 9]

Can permute the point ordering:

\[ \frac{||P_1 - P_3||}{||P_1 - P_2||} \cdot \frac{||P_4 - P_3||}{||P_4 - P_2||} \]

In the next slides we introduce a number of important definitions about lines and points in 2D and 3D and introduce the concepts of vanishing points and lines.

Lecture 4
Single View Metrology

Readings:
- [HZ] Chapter 2 "Projective Geometry and Transformation in 2D"
- [HZ] Chapter 3 "Projective Geometry and Transformation in 3D"
- [HZ] Chapter 8 "More Single View Geometry"
- [Hoiem & Savarese] Chapter 2

Silvio Savarese  Lecture 4 -  16-Jan-15
A line in 2D can be represented as the 3D vector $l = [a \ b \ c]^T$ in homogeneous coordinates. The ratio $-a/b$ captures the slope of the line and the ratio $-c/b$ defines the point of intersection of the line with the y axis. If a point $x$ belongs to a line $l$, then the dot product between $x$ and $l$ is equal to zero (Eq 10). This equation also defines a line in 2D.

In general, two lines $l$ and $l'$ intersect at a point $x$. This point is defined as the cross product between $l$ and $l'$ [Eq. 11]. This is easy to verify as the slide shows.

Proof: Given two intersecting lines $l$ and $l'$, the intersection point $x$ should lie on both lines $l$ and $l'$; thus the point $x$ is the intersection if and only if $x \cdot l = 0$ [Eq. 11] and $x \cdot l' = 0$ [Eq. 12]. Let $x$ be $l \times l'$. Then, the vector $x$ is perpendicular to the vector $l$ and the vector $l'$ and, thus, it satisfies the above constraints. Since the intersection is unique (set arbitrary $x' = x$ and show that $x'$ is $x$), $l \times l'$ is the point of intersection of the two lines.

Let now us compute the point of intersection of two parallel lines. We start by observing that a point at infinity in Euclidean coordinates corresponds to a point $x_{\infty}$ in homogeneous coordinates whose third coordinate is equal to zero.

Let us now consider two parallel lines $l$ and $l'$. When two lines are parallel, their slope is equal and thus $-a/b = -a'/b'$. Let’s now compute the point of intersection of these two lines. Using Eq. 11 we obtain Eq. 13 which is exactly the expression of a point at infinity (in homogenous coordinates). This confirms the intuition that two parallel lines intersect at infinity.

The point of intersection of two parallel lines returns a point at infinity which is also called ideal point.

* In Euclidean coordinates this point is at infinity
* Agree with the general idea of two lines intersecting at infinity
2D Points at infinity (ideal points)

\[ x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad x_3 \neq 0 \]

\[ l = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \]

Note: the line \( l = [a \ b \ c]^T \) pass trough the ideal point \( x_{\infty} \)

\[ l^T x_{\infty} = \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix} = 0 \quad \text{[Eq. 15]} \]

So does the line \( l' \) since \( a' b' = a b \)

One interesting property of a point at infinity is that all the parallel lines with the same slope \(-a/b\) passes through the point \([b \ -a \ 0]\) [Eq.15].

Lines infinity \( l_{\infty} \)

Set of ideal points lies on a line called the line at infinity. How does it look like?

\[ l_{\infty} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \]

Indeed:

\[ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix}^T \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0 \]

A line at infinity can be thought of the set of “directions” of lines in the plane.

Projective transformation of a point at infinity

\[ H = \begin{bmatrix} A & t \\ v & b \end{bmatrix} \]

\[ p' = H p \]

is it a point at infinity?

\[ H p_{\infty} = ? = \begin{bmatrix} A & t \\ v & b \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} p'_{x} \\ p'_{y} \\ p'_{z} \end{bmatrix} \quad \ldots \text{no!} \quad \text{[Eq. 17]} \]

\[ H a p_{\infty} = ? = \begin{bmatrix} A & t \\ 0 & b \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} p'_{x} \\ p'_{y} \\ 0 \end{bmatrix} \]

An affine transformation of a point at infinity is still a point at infinity.

We can further extend this concept and define the lines at infinity.

Consider 2 or more pairs of parallel lines (right side of the slide). Each pair of parallel lines intersect into a point at infinity. Let us call these set of points \( x_{\infty}, x_{\infty}' \ldots \). The line that passes through all these points at infinity must satisfy \( l^T x_{\infty} = 0, \ l^T x_{\infty}' = 0, \ldots \) and is simply \( l = [0 \ 0 \ c] \). Since \( c \) in an arbitrary value, we can simply write \( l = [0, 0, 1] \).

A line at infinity can be thought of the set of “directions” of lines in the plane.

Now, let’s see what happens if we apply a generic projective transformation \( H \) to a point at infinity \( p_{\infty} \) (Eq. 17).

Notice that the last element \( H p_{\infty} \) becomes non-zero which suggests that a projective transformation in general maps points at infinity to points that are no longer at infinity.

This is not true for affine transformations [Eq.18]. If we apply an affine transformation \( H \) to \( p_{\infty} \) will still obtain a point at infinity.
This time, let’s apply a projective transformation H to a line. The projective transformation of a line l is \( l' = H^{-T} l \) (Eq 19).

Let’s derive this equation. All points that pass through a line l must satisfy the line equation: \( x^T l = 0 \); \( \rightarrow x^T H^{-T} l = 0 \)

Since \( x' = H x \) and \( x'^T = x^T H^{-T} \), then \( x'^T H^{-T} l = 0 \); Because, after the transformation, a projected point \( x' \) must still belong to the projected line \( l' \) \( (x'^T l' = 0) \), it implies that \( l' = H^{-T} l \).

Let’s now apply the projective transformation \( \mathbf{H} \) to a line at infinity \( l_{\text{inf}} \). Is the projected line still at infinity? No.

Let’s now apply the affine transformation \( \mathbf{H}_{\mathbf{A}} \) to a line at infinity \( l_{\text{inf}} \). Is the projected line still at infinity? Yes, as the derivation next to Eq 21 shows.

Points and planes in 3D

Points in 3D (homogenous coordinates) are denoted as \( \mathbf{x} \) and their corresponding projective transformations in the 2D image are denoted as \( \mathbf{p} \).

Following the 2D case, we can represent a plane as a normal vector \((a,b,c)\) and a distance from the origin \(d\) which is \([\text{Eq.22}]\). Thus, a point is on a plane if and only if \([\text{Eq.23}]\) holds.

In 3D, it is tricky to represent a line but one can represent it using an intersection of 2 planes. See [HZ] Ch.3.2.2 for detail.

Philosophical interlude by C. Choy (course assistant): When you see an object, you are not perceiving 3D space (there is no way to 'sense' space directly), you are merely seeing the 2D perspective projection of the object on your retina. Thus seeing is equivalent to perspective projection. So it is essential to model the world as is in 3D space and perspective projection as we do.

In 3D, similarly to ideals points in 2D, points at infinity are defined as the point of intersection of parallel lines in 3D.
**Vanishing points**

The projective projection of a point at infinity into the image plane defines a vanishing point.

Similarly to the 2D case, by applying the projective transformation $M$ to a point at infinity $x_{\infty}$ we obtain a point $p_{\infty}$ in the image plane which is no longer at infinity.

Interestingly, similarly to the 2D case, the direction of the (parallel) lines in 3D associated to $x_{\infty}$ is given by the coordinates $x_1, x_2$ and $x_3$ of $x_{\infty}$.

Next we derive a useful relationship between parallel lines in 3D, the corresponding vanishing point in the image and the camera parameters.

Let's define as $d=[a, b, c]^T$ the direction of a set of parallel lines in 3D in the camera reference system. These lines intersect to a point at infinity and the projection of such point in the image returns the vanishing point $v$.

It's easy to prove that $v$ is related to $d$ via Eq. 24, where $K$ is the camera matrix. Equivalently, $d$ can be expressed as function of $v$ by Eq. 25, where the division by $\|K^{-1}v\|$ guarantees that $d$ has unit norm ($d$ is a direction).

The proof of Eq 24 is reported in the bottom part of the slide.

If we consider a plane $\pi$ as a superset of set of parallel lines, each set of parallel lines intersects at a point at infinity. The line that passes through such set of points at infinity is the line at infinity $l_{\infty}$ associated to $\pi$. A line at infinity is also defined as the line where two parallel planes intersect (in general, the intersection of two planes in 3D is a line). The projective transformation of $l_{\infty}$ to the image plane is no longer a line at infinity and is called the vanishing line or horizon line $l_{\text{hor}}$ (see Eq. 26). The horizon line is a line that passes through the corresponding vanishing points in the image.
Are these two lines parallel or not?

- Recognize the horizon line
- Measure if the 2 lines meet at the horizon
- If yes, these 2 lines are // in 3D

Recognition helps reconstruction!
Humans have learnt this

Vanishing points and planes

\[ \mathbf{n} = \mathbf{K}^\top \mathbf{l}_{\text{horiz}} \]

[Eq. 27]

It is easy to show that the normal \( \mathbf{n} \) of \( \Pi \) and \( l_{\text{horiz}} \) are related by Eq. 27 (see sec. 8.6.2 [HZ] for details), where \( \mathbf{K} \) is a camera matrix.

Again, Eq. 27 can be useful for estimating properties of the world. If we recognize the horizon and our camera is calibrated (\( \mathbf{K} \) is known), we can estimate the orientation of the ground plane.

Planes at infinity

- Parallel planes intersect at infinity in a common line — the line at infinity

Before introducing the last property that relates vanishing points and lines, we define the plane at infinity \( \Pi_\infty \).

A set of 2 or more vanishing lines (blue lines in the figure) defines the plane at infinity \( \Pi_\infty \) (yellow plane in the figure). The plane at infinity is described by \( [0 \ 0 \ 0 \ 1]^\top \) in homogenous coordinates.
Angle between 2 vanishing points

\[
\cos \theta = \frac{v_1^T \omega \ v_2}{\sqrt{v_1^T \omega \ v_1} \sqrt{v_2^T \omega \ v_2}} \quad \omega = (KK^T)^{-1}
\]

[Eq. 28]

If \( \theta = 90 \) \[v_1^T \omega \ v_2 = 0\] [Eq. 29]

Scalar equation

Projective transformation of \( \Omega_\infty \)

The matrix \( \omega \) has a special geometrical meaning in that it is the projective transformation of the absolute conic \( \Omega \) in the image. This relationship is expressed by [Eq. 30]. What's the absolute conic \( \Omega \)? It's a conic that lies in the plane at infinity \([0 0 0 1]^T\). Note that in general, the projective transformation of a quadric or conic \( \Omega \) is expressed as \( M^{-T} \Omega M^{-1} \) (see HZ page 73, eq. 3.16).

The matrix \( \omega \) satisfies other interesting properties:
- It depends only on the internal matrix and not on the extrinsic parameters of the camera \( R \) and \( T \).
- It is symmetric and known up to scale.
- If \( w_2 = 0 \), the camera has no skew.
- If \( w_1 = w_3 \), the camera has square pixels.

These properties are useful for two reasons:
- To calibrate the camera: By using Eq 28 or Eq 29, we can set up a system of equations that allows us to calibrate our camera — that is, to estimate internal parameters of the camera.
- To estimate the geometry of the 3D world: Once \( K \) is estimated or \( K \) is known, we can use these equations to estimate the orientation of planes in 3D w.r.t. to the camera reference system.

Let's see some examples next.

Why is this useful?

- To calibrate the camera
- To estimate the geometry of the 3D world
Lecture 4
Single View Metrology

- Review calibration
- Vanishing points and line
- Estimating geometry from a single image

Reading:
[H2] Chapter 2 "Projective Geometry and Transformation in 3D"
[H2] Chapter 3 "Projective Geometry and Transformation in 3D"
[H2] Chapter 8 “More Single View Geometry”
[Hoiem & Savarese] Chapter 2

Suppose we can identify two planes in an image of the 3D world (e.g., the two building facades) and suppose we can identify a pair of parallel lines on each of these planes. This allows to estimate two vanishing points in the image \( v_1 \) and \( v_2 \). Suppose we know that these planes are perpendicular in 3D (the two building facades are perpendicular in 3D). Than we can set up the system of equations above using Eq 29 and the definition of \( \omega \).

Is this sufficient to estimate the camera parameters? \( K \) has in general 5 degrees of freedom and Eq.29 is a scalar equation; so clearly we don’t have enough constraints.

Let’s now assume we can identify another pair of parallel lines and its corresponding vanishing point \( v_3 \). With a third vanishing point and by assuming that the 3 set of pairs of parallel lines are pairwise orthogonal in 3D (which is true in this example), we can use Eq. 29 to set up a system of 3 equations (constraints) (Eqs. 31).
Let's now also make some assumptions about the camera. Let's assume the camera is zero-skew and with square pixels.

Using the properties of $\omega$:
- $\omega$ is symmetric which means we have 6 unknowns.
- $\omega$ is known up to scale which reduces to 5 unknowns.
- $\omega_2 = 0$ and $\omega_3 = 0$, which reduces to 3 unknowns.

We have now enough constraints to solve for the unknowns and can calculate $\omega$. The actual parameters of $K$ can be computed from $\omega$ using the Cholesky factorization. We don't proof this results; for more details please refer to HZ pag 582.

Thus, at the end of this procedure we have managed to calibrate the camera from just one single image!

Once $K$ is known we can "reconstruct" the geometry of the scene; for instance, we can compute the orientation of all the planes in 3D using Eq. 27.

In order to do so, we need to identify the corresponding lines at infinity and select the orientation discontinuities (that is, where planes fold).

Again, the assortment of tools introduced in this lecture allows us to estimate properties of the camera from observations and/or estimate properties of the world by assuming we have some knowledge about the world (e.g., where the horizon lines are; planes discontinuities; etc...).

Notice that that actual scale of the scene cannot be recovered unless we assume we have access to some measurements in 3D (e.g., a window size); this is similar to what we did when we calibrated the camera.
Lecture 4 - Single View Metrology

- Review calibration
- Vanishing points and lines
- Estimating geometry from a single image
- Extensions

Reading:
- [HZ] Chapter 2 “Projective Geometry and Transformation in 3D”
- [HZ] Chapter 3 “Projective Geometry and Transformation in 3D”
- [HZ] Chapter 8 “More Single View Geometry”
- [Hoiem & Savarese] Chapter 2

Here we see some results obtained using the approach proposed by A. Criminisi and Zisserman (1999). This approach uses many of the results introduced in this lecture.

http://www.robots.ox.ac.uk/~vgg/projects/SingleView/models/merton/merton.wrl
La Trinita' (1426)
Firenze, Santa Maria Novella; by Masaccio (1401-1428)

http://www.robots.ox.ac.uk/~vgg/projects/SingleView/models/hut/hutme.wrl
A few years later, D Hoiem proposed an approach where the process of recovering the geometry from a single image is mostly automatic. This approach leverages recognition and segmentation results;
During the same period, A. Ng and his student A. Saxena at Stanford also demonstrated that it is possible to recover the geometry of the scene from a single image using recognition results and probabilistic inference.
In my own group, we have shown it is possible to combine recognition and reconstruction from a single image in a coherent formulation.

Next lecture:

Multi-view geometry (epipolar geometry)

Appendix